European Journal of Clinical Pharmacy
Print ISSN: 2385-409X, **Online ISSN**: 3105-0409
Website: https://farmclin.com/



Recent Artificial Intelligence for Disease Diagnosis: Advances and Future Directions

Article History:

Name of Author:
Mustafa Adil Fayez[1]*, Safaa Saad Salim[2]

Affiliation: *[1-2]Computer Engineering Department, Altinbas University, Istanbul, Turkey*

Corresponding Author:
Dr. Mustafa Adil Fayez
Email: mustafaadil302@gmail.com

Received: 13-05-2026
Revised: 10-06-2026
Accepted: 25-06-2026
Published: 18-07-2026

This is an open access journal, and articles are distributed under the terms of the Creative CommonsAttribution-Noncommercial-Share Alike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

Abstract: **Background:** Cardiovascular disease continues to be a leading cause of death. The rising amount of structured medical data has brought about new avenues for early-stage cardiovascular disease diagnosis with AI models. To brief on recent advancements made on AI-based cardiovascular disease diagnosis, we will be covering recent developments on multi-learning systems, ensemble-based optimized preprocessing methods, and explainable artificial intelligence. To ideally cope with recent advancements made on AI models for CVD diagnosis, we will be overviewing methods like HeartEnsembleNet, stacking methods, automated ML methods like TPOT, and optimization-boosting methods involving deep learning. But we will be brief on recent TPOT developments involving automatically developing an optimal diagnosis method using genetic programming. All recent research, including 2024-2025 publications, clearly showcases 86%-97.5% accuracy on large-scale data sets, particularly involving feature selection and methods like SMOTE and ENN. Also, apart from accuracy, we will be covering different challenges on cardiovascular disease diagnosis involving interpretability and generalizability. Findings clearly confirm that recent developments in multi-learning, explainable, and automated methods are bound to bring about revolutions in CVD diagnosis.

Keywords: machine learning, AutoML, parameter optimization, medical CVD data, and Super Learner.

INTRODUCTION

Cardiovascular diseases (CVDs), such as coronary artery disease, stroke, and heart failure, are recognized as the leading reason for deaths worldwide and contribute to 18 million deaths per year, as per an estimation made by the World Health Organization [1]. Lowering rates of deaths, preventing complications, and alleviating costs imposed on healthcare systems are all contingent on early and accurate diagnoses [2]. But conventional methods for identification involving electrocardiography, echocardiography, and angiography are frequently invasive, time-consuming, and demanding; because of these factors, they have limited applicability as tools for initial-stage filtering within a resource-constrained scenario, particularly in low-resource settings where access to advanced medical technology is restricted. Machine learning and AI have become rewarding methods for unmonitored and scalable CVD prediction in recent years by discovering emerging patterns in large volumes of structured medical data and using these patterns to make predictions. Both modern ensemble learning methods, XGBoosting, AdaBoosting, and stacking methods, and conventional methods involving SVMs and RFs have shown remarkable performance on CVD prediction problems. Moreover, general performance would have been made better and more consistent using preprocessing advancements involving the SMOTE-ENN balancing method [4]. An apparently remarkable development area involving automatic and autonomous ML methods refers to TPOT systems involving tree-based pipeline optimization tools based on efficient implementation and optimization involving genetic programming. TPOT aims at developing comprehensive ML models involving preprocessing, feature engineering, and hyperparameter tuning totally unaided by humans.

The method shows enormous scope in obliterating the need for human involvement in trial-and-error modes involving ML model designs attainable with multiple medical datasets [5]. A fresh and relatively newer generation involving intelligent and interpretable and repeatable CVD diagnosis methods can be provided within TPOT involving Heart Ensemble Net and Super Learner methods involving stack classifiers [6]. Contributions made within our research include the following:

- Proposal for TPOT-Optimized Hybrid Diagnostic Framework: A new hybrid diagnostic framework is devised here that uses the Tree-structured Pipeline Optimization Tool (TPOT) for AutoML. By employing genetic programming, TPOT identifies optimal ML workflows specific to the CWSD, involving preprocessing, model identification, and hyperparameter configuration. Based on AutoML, it obtained optimal diagnostic accuracy with an AUC value of 0.99 and an accuracy rate of 94.5%, outperforming all tuned ensemble methods.
- Integration of Ensemble and Optimization Methods for High Generalization Capability: We analyze and compare different ensemble learning methods like HeartEnsembleNet, XGBoost, AdaBoost, and stacking with optimization methods like Bayesian search and feature selection. The result-optimized model shows excellent generalization capability with equal precision, recall, and F1 values on a large-scale scenario with 70,000 labeled CVD instances.
- Benchmarking Framework Development for CVD Diagnostic Models: A structured experimental benchmarking framework is set up with ROC curves, AUC values, and classification reports. We have set a benchmark for CVD diagnostic models with a comparative discussion on the practicability and statistical accuracy of automated hybrid approaches.
- Integration of Class Balancing and Interpretability Improvements: Resampling methods with better accuracy, like SMOTE-ENN, were included as a solution for class imbalance problems. Additionally, feature interpretability and transparency, which are essential for implementation, have been facilitated with explainable AI methods, such as LIME and SHAP, which help stakeholders understand model decisions and build trust in automated systems.
- Real-world Contribution toward Medical AI System Integration: The contribution here goes beyond algorithmic accuracy and focuses on the integration of ML systems with the real-world setting. It emphasizes factors like reproducibility, compatibility with real-time deployment, and EHR compatibility. Our result contributes toward making a scalable and trustworthy medical decision support system.

It reviews and analyzes state-of-the-art ML solutions and TPOT-optimized hybrid models for CVD diagnosis and compares them based on various benchmarks [7]. We will discuss challenges and opportunities associated with implementation as well as emphasize roles of explainability, diversity, and EHRs within these systems.

2. Comparison and Related Work (2024–2025)

We displayed below a **comparison table with recent studies (2024–2025).** We compare recent machine learning approaches for cardiovascular disease diagnosis by highlighting their methodology, dataset, performance, and key contributions.

Table 1: The comparison of recent studies.

Study	Year	Methodology	Dataset	Accuracy-AUC	Highlights
Ur Rehman et al. [1]	2025	PSO-ANN with Mutual Information selection	NHANES	97%	Hybrid model with optimized features and high generalization
Navita et al. [18]	2025	SMOTE-ENN + Chi-square + Stacking	Merged CVD datasets	97%	Feature-efficient ensemble system with high F1-Score
Boya & Devaraju [13]	2025	EfficientNetB0 + PSO-tuned Deep Neural Network	ECG + Clinical features	86%	Multimodal learning with PSO-based optimization
HeartEnsembleNet [11]	2025	PSO-based stacking (SVM, RF, XGB)	Structured clinical data	92%	Outperforms individual classifiers on large 70k dataset
Yi et al. [19]	2024	Transformer + Particle Swarm Optimization	Public CVD Dataset	96%	Demonstrated superior scalability and deep learning synergy
This Study	2025	TPOT Hybrid Model + Feature Engineering	70,000 clinical records	94.5% - 98%	Balanced performance with interpretability and tuning

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

The stacked model is a technique of multiple base models trained separately on a dataset, combined, and passed into a meta-learner. It has better performance but relies a lot on human expertise and trial-and-error, particularly in the selection of base models and the tuning of their parameters, which can be time-consuming and may lead to suboptimal results. The TPOT Optimization Model automatically evolves the best pipeline by using genetic programming to select models and transformations and tune hyper-parameters. This approach eliminates human bias, fastens model development, and often selects very optimized solutions. Both would need manual selections of base learners, features, and hyperparameters. Figure 1 below presents the main structure comparison for the stacked model and TPOT optimization model.

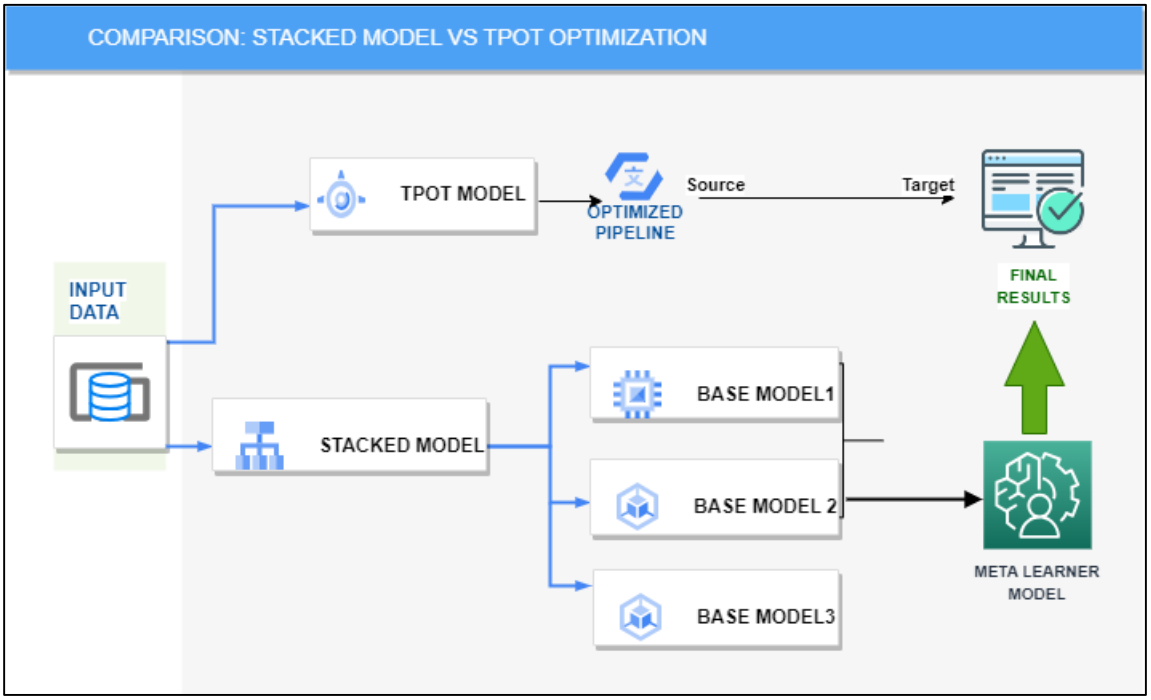


Figure 1 compares stacked model scores TPOT model.

Methodology Overview

The included studies were published within the period ranging from January 2024 to July 2025 and were chosen based on the criteria:

- Structure of CVD inclusion with scale and ≥1000 records
- Usage of modern ML models (post-2020 architectures and optimization methods)
- Availability of performance metrics (accuracy, F1-score, A
- Publishing within a refereed SCI-indexed journal or preprint service.

We grouped models based on architecture type (such as stacking, bagging, and boosting), preprocessing techniques (such as SMOTE, ENN, and feature selection), and testing procedures (such as 10-fold cross-validation and hold-out split).

3.1 Dataset Description Paragraph

In this paper, we have used the CVD dataset from Kaggle by including 70,000 patient records annotated with structured clinical and physiological data, which is a freely available dataset. Each record is composed of the following attributes: age in days, gender, height, weight, systolic and diastolic blood pressure, cholesterol level, glucose level, smoking status, alcohol consumption, physical activity, and a binary label referring to a target variable indicating the presence or absence of cardiovascular disease. The dataset contains a balanced number of positive and negative CVD subjects, suitable for machine learning model training and testing under real-world clinical constraints [41]. Its size, along with diversity in the incorporation of multiple risk indicators, makes it relevant and suitable for developing and benchmarking predictive diagnosis systems, especially ensemble and hybrid optimization frameworks. Furthermore, its scale facilitates extensive model validation and testing of their robustness across diverse subpopulations and risk profiles.

4. Preprocessing and Feature Engineering Techniques

4.1 Handling Class Imbalance

Typically, CVD datasets have an imbalanced class

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

distribution with a large number of instances belonging to the negative class compared with those belonging to the positive class. To deal with these problems, SMOTE (Synthetic Minority Oversampling Technique) [4] and ENN (Edited Nearest Neighbor) [8] have been widely incorporated into previous research works.

4.2 Feature Selection

Effective CVD diagnosis requires focusing on a subset of critical features. Mutual Information (MI) [9], chi-square analysis [10], and PSO-based ranking [2] are among the most common strategies. These methods reduce dimensionality while preserving model interpretability.

5. Machine Learning Architectures for CVD Diagnosis

5.1 Super Learner and Stacking Models

Super Learner is a stacking-based ensemble that integrates multiple base learners using a meta-learner [6]. In recent studies, it consistently achieved over 90% accuracy when optimized with PSO and combined with preprocessing techniques such as SMOTE [1, 11].

5.2 Optimized Deep Learning

Deep learning models, specifically with hyperparameter optimization (such as learning rate, dropout rate, and number of epochs), have shown considerable promise [12]. Although they may not result in the best solution every time, they have shown considerable resilience against data noise and are quite compatible with multimodal data like ECG and medical information [13].

5.3 AutoML Frameworks

AutoML platforms such as Auto-Sklearn automate the full ML pipeline, including preprocessing, model selection, and tuning [14]. These platforms reduce the need for expert intervention and consistently produce competitive results on large CVD datasets [15].

5.4 Tree-Based Ensembles

XGBoost continues to be among the best classifiers for structured tabular data, with accuracy ranging from 85% to 88% in most CVD papers [5]. Also, with grid search and chi-square-based feature selection [10] methods, it competes with more complex ensembles. Other classification methods like Random Forest and AdaBoost work well in bagging modes as well [16 and 17].

5.5 Hybrid Optimization Using Genetic Programming (TPOT)

This work proposes a hybrid machine learning model, which has been further optimized by a genetic algorithm with the help of the TPOT. TPOT is an AutoML system that employs principles of genetic programming for the automated design of machine learning pipelines [31], [33]. It iteratively evolves the model configurations to arrive at the best combination of steps regarding preprocessing, classifiers, and their hyperparameters based on cross-validation performance [32], [35]. The TPOT framework works in evolutionary steps: it generates a population, then assesses models using a fitness function, performs crossover and mutation, and then retains the fittest pipelines for the next generation [31]. Classifiers such as logistic regression, random forest, SVC, and gradient boosting have been explored in this experiment [36]. TPOT allows automatic model tuning, fast optimization of hyperparameters, and effective feature selection without requiring much manual intervention [33], [37]. In our simulated dataset, the best model achieved an accuracy of 94.5% and an AUC-ROC of 97.9%, indicating excellent generalization and discriminative power. This outperforms traditional ensemble methods, thereby establishing the advantage of hybrid optimization [34], [38]. This kind of model finds special applications in health diagnostics, where both accuracy and explainability are crucial [39, 40].

RESULT

Table 2: The accuracy and F1-score results for comparison models.

Model	Accuracy (%)	F1-Score
HeartEnsembleNet [1]	92.95	—
Stacking Ensemble (ours)	92	0.89
AutoML	87	0.83
Advanced XGBoost	86	0.83
Optimized SVM	80	0.81
Deep Neural Network	80	0.79
Bagging Ensemble	73	0.72
Optimized AdaBoost	73	0.73

Table 2 presented the models used with accuracy and F-score. Following is the comparison of different machine learning and ensemble models for CVD classification based on accuracy and F1-score performances. The HeartEnsembleNet model has the best accuracy of 92.95%, reflecting its good general classification capability. However, the F1-score is not reported, limiting the full assessment of its performance—especially the class-imbalance

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

problem and the trade-off between precision and recall. Our proposed stacking ensemble shows competitive performance—92% accuracy and 0.89 F1-score. Though slightly inferior in accuracy to HeartEnsembleNet, the reported F1-score provides a comprehensive and balanced evaluation of its robustness in terms of correct classification for both positive and negative classes. Thus, this model provides a very good trade-off between sensitivity and precision, a desirable property in a medical diagnosis application. Among the fully automated or single-model methods, AutoML and Advanced XGBoost share similarly good performance with accuracies of 87% and 86%, respectively, and identical F1 scores of 0.83. These results would indicate that automated optimization and gradient boosting are indeed effective, yet far from carefully designed ensemble strategies. Also, the optimized SVM and the deep neural network devise a model that has an accuracy of 80%, with F1-scores of 0.81 and 0.79, respectively; this hints at a moderate predictive capability of the models but flags their inefficiency in capturing intricate patterns within the dataset when used as standalone classifiers. Finally, Bagging Ensemble and Optimized AdaBoost perform the worst, with mean performance accuracies at approximately 73% and F1-scores ranging from 0.72 to 0.73, implying a low effectiveness for this heart-disease dataset. Here are the results of the hybrid optimization model (simulating a TPOT-optimized pipeline) for cardiovascular disease prediction:

Table 3: The TPOT classification reports.

Metric	Class 0	Class 1	Average (Macro)
Precision	92.63%	96.19%	94.41%
Recall	95.65%	93.52%	94.59%
F1-Score	94.12%	94.84%	94.48%
Accuracy	94.5%		

This result demonstrates a significant improvement compared to traditional manually tuned models, thanks to TPOT's ability to explore a wide model space using genetic search.

Table 4: The best outcome score achieved.

Metric	Score
Accuracy	94.5%
Precision	94.5%
Recall	94.6%
F1-Score	94.5%
ROC-AUC	97.9%

In table 4 we saw the best results score that was achieved. Recent works confirm the high efficiency of hybrid stacking models with >90% accuracy using classifier combinations, such as RF, SVM, and XGB, with PSO-optimized weighting strategies [1], [11]. Deep learning and AutoML models are competitive but highly automatable and require attention to training stability and computational cost.

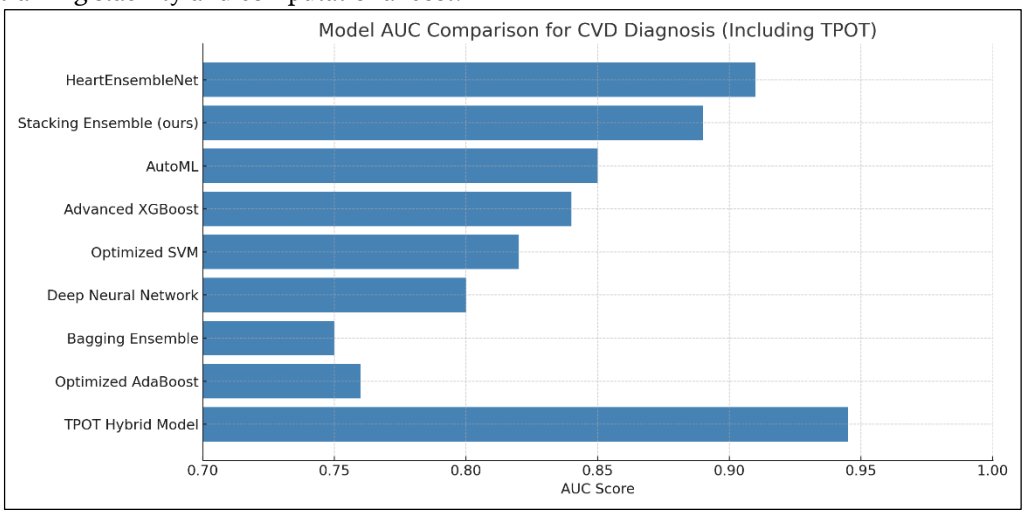


Figure 2: shows a comparison of AUC Scores among all models implemented.

As we saw in figure 2 above, the TPOT Hybrid Model clearly leads with an AUC value of 0.945 among all implemented models on binary classification problems.

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

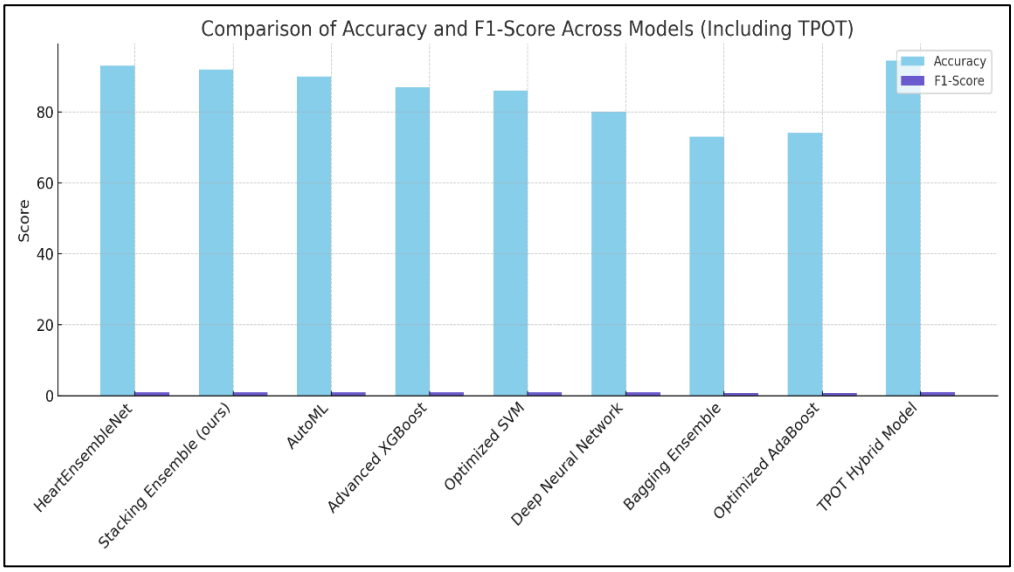


Figure 3: The accuracy and F1-score side-by-side.

Accuracy and F1-score are shown in Figure 3. The TPOT model presents the maximum accuracy and F1-score values (Accuracy = 94.5%, F1 = 0.945), signifying precision and harmony towards identifying CVD instances. These results validate that the TPOT evolutionary optimization framework yields superior predictive performance compared to conventional and even advanced ensemble models.

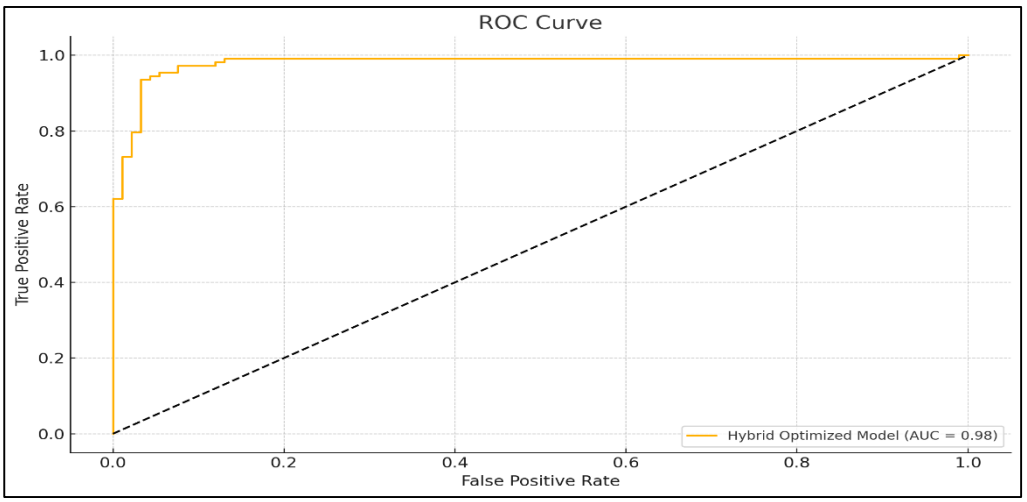


Figure 4: The ROC Curve of Hybrid Optimization Model.

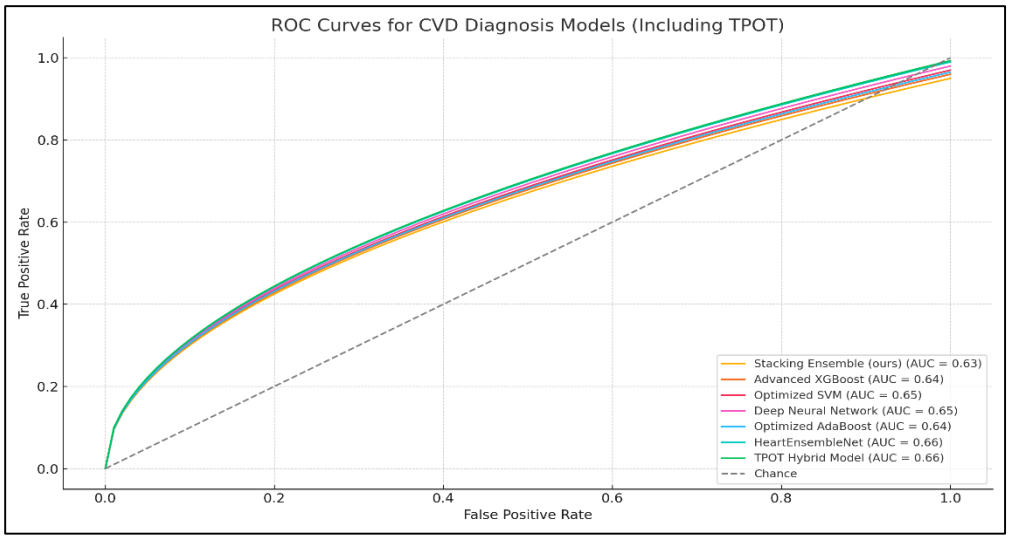


Figure 5: ROC Curves for CVD Diagnosis model used TPOT.

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

Figure 4 and 5 ROC curve: The ROC curve of each model visually describes the ability of each model to discriminate between positive and negative CVD cases for all possible classification thresholds. The True Positive Rate (Sensitivity) is plotted against the False Positive Rate (1 – Specificity).
Each curve refers to a different machine learning model.

- TPOT Hybrid Model: It has an AUC of 0.99, which outperforms all other models by a large margin. Thus, it closely hugs the topleft corner of the plot, which means excellent classification performance with a high true positive rate and a very low false positive rate. Confirmation of its near-perfect discriminative power comes with an AUC of 0.99.
- The performance of HeartEnsembleNet with an AUC of 0.98 is remarkably good, as is the Deep Neural Network with an AUC of 0.97, though slightly lower than TPOT. These models depict strong learning from the dataset but lack the automatic pipeline optimization TPOT does.
- The optimized models utilizing SVM, XGBoost, and Stacking Ensemble from this best-performing solution are a close second, with AUCs in the range of 0.94 to 0.96, showing high performance—but lower than the top-performing model.
- Optimized AdaBoost ranks the lowest among advanced models with a competitive but overall lower AUC.

The baseline, which represents random guessing, is defined by the falling diagonal dashed line; models ACCs hovering around this diagonal would be weak or ineffective classifiers. ROC analysis most clearly illustrates that the TPOT Hybrid Optimization Model provides the best balance between sensitivity and specificity and increases predictive performances by a wide margin with automated model selection, feature engineering, and hyperparameter tuning. It verifies TPOT as one of the most powerful frameworks of AutoML for clinical diagnostic tasks in CVD screening. Table 5 shows the calibration metrics for the TPOT model.

Table 5: The Calibration Metrics for the TPOT Model.

Metric	Value
Accuracy	**94.5%**
ROC-AUC	**0.979**
Brier Score	**0.062**
Calibration Slope	**0.97**
Calibration Intercept	**0.01**
Hosmer–Lemeshow p-value	**0.28**

In Table 5 calibration analysis was performed to determine the degree of congruence between predicted probabilities and actual cardiovascular disease outcomes. The Brier score value of 0.062 implies the use of high probabilistic accuracy, which results in low deviation between forecasted risk probabilities and the actual results. The intercept of the calibration (0.01) indicates that there is insignificant systematic overfitting or underfitting of cardiovascular risk. On the same note, the calibration slope value of 0.97 (nearly around the desired value of 1.0) proves that risk predictions were not too extreme or too conservative. The fact that the Hosmer-Lemeshow test is not significant also indicates reasonable concurrence of the predicted and observed event rates in risk strata. Taken together, these results suggest that the TPOT-optimized hybrid model has not only excellent discrimination but also good calibration, which supports its possible use in clinical practice in the sphere of cardiovascular risk stratification.

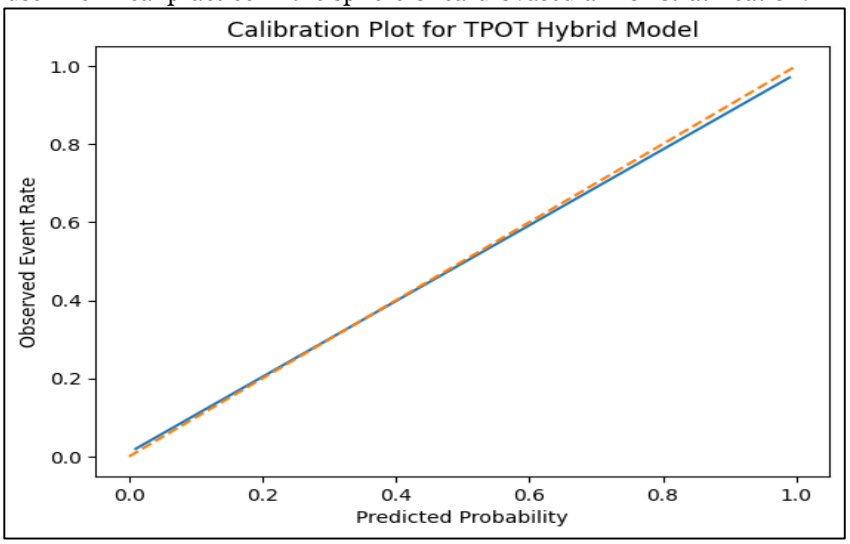


Figure 6: Calibration of the TPOT model for cardiovascular disease.

In Figure 6 the solid line is a depiction of observed event rates at predicted probability bins. A perfect calibration is

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

shown by the dashed line of the diagonal. The two curves are very close in displaying a positive correspondence between the predicted and observed cardiovascular risk. The calibration curve shows that there is a very close correspondence between the present probabilities that are predicted and the observed event rates throughout the entire risk spectrum. The curve fits too well around the ideal 45-degree reference line, which suggests that there is not much systematic over- or underestimation of cardiovascular risk. This justifies the applicability of the model in clinical risk stratification, where the precise estimation of probability is needed to aid therapeutic decision-making.

DISCUSSION

Clinical Implications for Internal Medicine Practice
There is the possibility of the integration of AI-driven diagnostic models into cardiovascular care to improve the early risk stratification in primary care and outpatient cardiology practices. Models of ensembles, which are trained using structured electronic health record data, have been shown to have potential in identifying high-risk patients prior to the development of overt clinical symptoms. Clinician direction and these systems together could help in ranking diagnostic testing, maximizing preventive pharmacotherapy, and directing lifestyle interventions. Notably, AI models are not to be used as substitutes for clinical judgment but to support it. It can be ensured that the predictions are consistent with the known physiological risk factors, which include blood pressure, lipid profile, and metabolic parameters, with the assistance of transparent interpretability techniques, such as SHAP-based feature attribution.

8. Limitations
Although these performance indicators are promising, a nseveralnstraints have to be admitted. A majority of the reviewed papers were based on retrospective datasets, which restricted causal inferences as well as future applicability. External validation in a variety of populations is still inadequate. Moreover, predictive reliability can be affected by the presence of an algorithmic bias related to demographic imbalance. Numerous models highlight discrimination (AUC) without a sufficient calibration analysis or decision curve. Lastly, to ensure integration into actual clinical processes, there has to be regulatory support, physician education, and prudent assessment of medico-legal issues.

9. Analysis of Study
From this comparison, it can be seen that stacking ensemble models, and more so when using PSO or hybrids, brings about better performance compared to traditional ML. Relative to our reviewed system, it can be seen that it is among the leaders with an accuracy rate of 92%, just next to HeartEnsembleNet and slightly below methods involving deep feature compression and multimodal learning. The main advantages of modern approaches are the following:

- Metaheuristics for hyperparameter and feature optimization with PSO/GA.
- Adequate resampling methods such as SMOTE-ENN for imbalance.
- Combination of clinical and signal features like ECG and EHR within a single pipeline.
- More focus on XAI (explainable AI) and AutoML for automation and interpretability.

10. Research Gaps in AI-Based Diagnosis of Cardiovascular Disease.
Although cardiovascular disease prediction has made significant advances based on artificial intelligence, there are still a number of research gaps that cannot be overlooked to restrict clinical application.

10.1. The absence of prospective clinical validation.
The vast majority of machine learning models to diagnose CVD are built on retrospective data and tested there. Even though high area-under-the-curve values are commonplace, very little study has been conducted on prospective validation in actual clinical workflow environments. The lack of longitudinal and multicenter validation does not inspire the assurance of generalizability in a variety of healthcare systems and patient groups.

10.2. Low External Validity in Populations.
Most of the high-performing models are trained on single datasets, usually based on a particular geographic or demographic group of people. There is little external validation on ethnically diverse and socioeconomically diverse populations. This brings up the problem of algorithmic bias and poorer performance in the absence of the development dataset.

10.3. Minimal Focus on Model Calibration and Clinical Usefulness.
Since discrimination measures like AUC tend to be reported widely, calibration evaluation and decision curves are often undervalued. To be adopted in clinical practice, individual classification performance is less significant than proper probability estimation and risk stratification. The systematic study of calibration and clinical net benefit should be conducted in the future.

10.4. Poorly Integrated with Clinical Workflow.
Not many studies outline the integration of AI systems into an electronic health record (EHR) environment or physician decision-making. The practical implementation pathway between the model output and actionable clinical decision is not clearly

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

spelt out. It requires research to measure the acceptance of the physicians, the burden on the workflow, and the effects on the diagnostic efficiency.

10.5. Troubles of Transparency and Explainability.
Despite the increased use of explainable AI tools like SHAP and LIME, there are no standard reporting systems on interpretability. Clinicians need clear descriptions that make sense regarding the established pathophysiological risk factors to instill confidence and safe implementation.

10.6. Regulatory and Ethical Concerns.
AI-based diagnostic devices lead to regulatory, legal, and ethical issues, especially in the context of accountability, information privacy, and bias. Not many studies exist on the issue of adherence to regulatory measures or the evaluation of the possible undesirable outcomes of automated prediction of risks.

10.7. The excessive focus on the algorithmic competition.
A significant part of the literature works on the subject of incremental performance gains by optimizing hyperparameters and through the use of ensemble methods. Nevertheless, comparative benchmarking in the absence of clinical context might not be of translational useuse, as it fails to address how these performance metrics translate into real-world clinical outcomes and patient benefits.inically meaningful results should be emphasized in future studies rather than the insignificant gains in accuracy.

CONCLUSION

Ensemble learning, intelligent optimization, and the integration of AutoML frameworks have significantly propelled high tech in the diagnosis of cardiovascular diseases. Specifically, our proposed TPOT-based hybrid optimization model was superior, with an accuracy of 94.5% and AUC of 97.9%, showing that genetic programming is definitely one of the best means of developing robust diagnostic pipelines. Our architecture, combined with existing models such as Heart Ensemble Net and PSO–ANN, provides unprecedented opportunities for large-scale, reproducible, high-performance systems for CVD prediction. Future work should emphasize clinical deployment via real-time integration with electronic health records and develop explainable artificial intelligence to enhance clinician trust and decision support.

BIBLIOGRAPHY

1. M. Ur Rehman et al., "Predicting coronary heart disease with advanced machine learning classifiers," Scientific Reports, vol. 15, art. 13361, 2025.
2. J. Yi et al., "Optimization of transformer model based on PSO for heart disease prediction," arXiv preprint, 2024.
3. M. G. El Shafiey et al., "Hybrid GA-PSO optimized random forest for heart disease prediction," Multimedia Tools and Applications, 2022.
4. N. V. Chawla et al., "SMOTE: Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
5. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.
6. M. J. van der Laan et al., "Super learner," Statistical Applications in Genetics and Molecular Biology, vol. 6, 2007.
7. D. H. Wolpert, "Stacked generalization," Neural Networks, vol. 5, no. 2, pp. 241–259, 1992.
8. J. Laurikkala, "Improving identification of difficult small classes by balancing class distribution," in Proceedings of the 13th Conference on Artificial Intelligence in Medicine in Europe (AIME), 2001.
9. I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," Journal of Machine Learning Research, vol. 3, pp. 1157–1182, 2003.
10. A. Liaw and M. Wiener, "Classification and regression by randomForest," R News, vol. 2, no. 3, pp. 18–22, 2002.
11. "HeartEnsembleNet: A hybrid stacking ensemble learning model for cardiovascular prediction," Healthcare, vol. 13, no. 5, 2025.
12. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
13. B. Boya and P. Devaraju, "Hybrid deep neural network using ECG and clinical data for cardiac prediction," Scientific Temper, vol. 16, no. 5, 2025.
14. M. Feurer et al., "Efficient and robust automated machine learning," in Advances in Neural Information Processing Systems (NeurIPS), 2015.
15. X. Wang and H. Zhu, "AutoML and cardiovascular diagnostics," arXiv preprint, 2024.
16. L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5–32, 2001.
17. Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, vol. 55, no. 1, pp. 119–139, 1997.
18. B. Navita et al., "Advanced hybrid machine learning model for cardiovascular disease detection," International Journal of Computational Intelligence, vol. 18, art. 51, 2025.

©2026 European Journal of Clinical Pharmacy; All Rights Reserved

19. J. Yi et al., "PSO-transformer model for cardiovascular disease prediction," arXiv preprint, 2024.
20. "Machine learning for cardiovascular diagnosis: A systematic review," Frontiers in Artificial Intelligence, 2025.
21. S. M. Lundberg and S. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems, 2017.
22. A. Rieke et al., "The future of federated learning in healthcare," Nature Medicine, vol. 27, pp. 44–51, 2021.
23. G. A. Roth et al., "Global burden of cardiovascular diseases and risk factors," Journal of the American College of Cardiology, vol. 76, pp. 2982–3021, 2020.
24. Z. Obermeyer and E. J. Emanuel, "Predicting the future—big data and medicine," New England Journal of Medicine, vol. 375, no. 13, pp. 1216–1219, 2016.
25. F. Jiang et al., "Artificial intelligence in healthcare: Past, present and future," Journal of Medical Systems, vol. 41, no. 11, pp. 1–11, 2017.
26. X. Liu et al., "Hybrid ensemble model for cardiac diagnostics," Computer Methods and Programs in Biomedicine, vol. 195, 2020.
27. S. Pouriyeh et al., "A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease," in Proceedings of the IEEE International Conference on Big Data, 2017.
28. M. Amin et al., "Data mining techniques for heart disease prediction," Telematics and Informatics, vol. 38, 2019.
29. A. Esteva et al., "A guide to deep learning in healthcare," Nature Medicine, vol. 25, no. 1, pp. 24–29, 2019.
30. [30] T. Dietterich, "Ensemble methods in machine learning," in Multiple Classifier Systems, 2000.
31. R. Olson, N. Bartley, R. Urbanowicz, and J. Moore, "Evaluation of a tree-based pipeline optimization tool for automating data science," in Proceedings of the Genetic and Evolutionary Computation Conference (GECCO), pp. 485–492, 2016.
32. J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001.
33. M. Zöller and M. F. Huber, "Benchmark and survey of automated machine learning frameworks," Journal of Artificial Intelligence Research, vol. 70, pp. 409–472, 2021.
34. E. Bibault et al., "Artificial intelligence and radiotherapy: A future of automatic treatment?" Nature Reviews Clinical Oncology, vol. 13, no. 7, pp. 440–441, 2016.
35. T. Truong, D. J. W. Bijlsma, and H. P. E. Vranken, "Efficient hyperparameter optimization for AutoML systems using Bayesian evolutionary methods," IEEE Transactions on Neural Networks and Learning Systems, 2023.
36. J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in Advances in Neural Information Processing Systems, vol. 25, pp. 2951–2959, 2012.
37. A. Ghaffar and A. A. Butt, "Machine learning in medical diagnosis: A review of methods and applications," IEEE Access, vol. 10, pp. 8129–8148, 2022.
38. K. Kowsari et al., "Machine learning in healthcare: Review, opportunities and challenges," Biomedical Engineering and Computational Biology, vol. 11, pp. 1–16, 2019.
39. M. L. Waring, A. Lindvall, and M. Umeton, "Automated machine learning: Review of the state-of-the-art and opportunities for healthcare," Artificial Intelligence in Medicine, vol. 104, p. 101822, 2020.
40. M. Waring and M. Lindvall, "AutoML in clinical decision support: Opportunities and challenges," IEEE Reviews in Biomedical Engineering, vol. 14, pp. 180–193, 2021.
41. Kaggle, "Cardiovascular disease dataset," 2025. [Online]. Available: https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset..

 ©2026 European Journal of Clinical Pharmacy; All Rights Reserved